

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/785,632A

DATE: 05/25/2001

TIME: 19:48:04

Input Set : A:\Pto.amc

Output Set: C:\CRF3\05252001\I785632A.raw

P. 5

4 <110> APPLICANT: Kim, Jin-Soo  
5 Kwon, Young Do  
6 Kim, Hyun-Won  
7 Ryu, Eun-Hyun  
8 Hwang, Moon-Sun  
10 <120> TITLE OF INVENTION: ZINC FINGER DOMAINS AND METHODS OF  
11 IDENTIFYING SAME  
13 <130> FILE REFERENCE: 12279-002001  
15 <140> CURRENT APPLICATION NUMBER: 09/785,632A  
16 <141> CURRENT FILING DATE: 2001-02-16  
18 <160> NUMBER OF SEQ ID NOS: 166  
20 <170> SOFTWARE: FastSEQ for Windows Version 4.0  
22 <210> SEQ ID NO: 1  
23 <211> LENGTH: 10  
24 <212> TYPE: DNA  
25 <213> ORGANISM: HIV-1  
27 <400> SEQUENCE: 1  
28 gacatcgagc 10  
30 <210> SEQ ID NO: 2  
31 <211> LENGTH: 10  
32 <212> TYPE: DNA  
33 <213> ORGANISM: HIV-1  
35 <400> SEQUENCE: 2  
36 gcagctgctt 10  
38 <210> SEQ ID NO: 3  
39 <211> LENGTH: 10  
40 <212> TYPE: DNA  
41 <213> ORGANISM: HIV-1  
43 <400> SEQUENCE: 3  
44 gctggggact 10  
46 <210> SEQ ID NO: 4  
47 <211> LENGTH: 10  
48 <212> TYPE: DNA  
49 <213> ORGANISM: Homo sapiens  
51 <400> SEQUENCE: 4  
52 aggggtggagt 10  
54 <210> SEQ ID NO: 5  
55 <211> LENGTH: 10  
56 <212> TYPE: DNA  
57 <213> ORGANISM: Homo sapiens  
59 <400> SEQUENCE: 5  
60 gctgagacat 10  
62 <210> SEQ ID NO: 6  
63 <211> LENGTH: 47  
64 <212> TYPE: DNA  
65 <213> ORGANISM: Artificial Sequence  
67 <220> FEATURE:

## RAW SEQUENCE LISTING

DATE: 05/25/2001

PATENT APPLICATION: US/09/785,632A

TIME: 19:48:04

Input Set : A:\Pto.amc

Output Set: C:\CRF3\05252001\I785632A.raw

```

68 <223> OTHER INFORMATION: optimal binding site
70 <400> SEQUENCE: 6
71 ccggcgtggg cggctgcgtg ggcgtgcgtg ggcggactgc gtgggcg      47
73 <210> SEQ ID NO: 7
74 <211> LENGTH: 47
75 <212> TYPE: DNA
76 <213> ORGANISM: Artificial Sequence
78 <220> FEATURE:
79 <223> OTHER INFORMATION: optimal binding site
81 <400> SEQUENCE: 7
82 tcgacgcca cgcagtccgc ccacgcacgc ccacgcagcc gcccacg      47
84 <210> SEQ ID NO: 8
85 <211> LENGTH: 49
86 <212> TYPE: DNA
87 <213> ORGANISM: HIV-1
89 <400> SEQUENCE: 8
90 ccggcgagcg ggcggtcgag cgggcgtgag cgggcggatc gagcgggcg      49
92 <210> SEQ ID NO: 9
93 <211> LENGTH: 49
94 <212> TYPE: DNA
95 <213> ORGANISM: HIV-1
97 <400> SEQUENCE: 9
98 tcgacgccc ctcgatccgc ccgctcacgc ccgctcgacc gcccgctcg      49
100 <210> SEQ ID NO: 10
101 <211> LENGTH: 50
102 <212> TYPE: DNA
103 <213> ORGANISM: HIV-1
105 <400> SEQUENCE: 10
106 ccggctgctt gggcggctgc ttgggcgtgc ttgggcgggc tgcttgggcg      50
108 <210> SEQ ID NO: 11
109 <211> LENGTH: 50
110 <212> TYPE: DNA
111 <213> ORGANISM: HIV-1
113 <400> SEQUENCE: 11
114 tcgacgcca agcagcccgc ccaagcacgc ccaagcagcc gcccagcag      50
116 <210> SEQ ID NO: 12
117 <211> LENGTH: 47
118 <212> TYPE: DNA
119 <213> ORGANISM: HIV-1
121 <400> SEQUENCE: 12
122 ccggactggg cgggggactg ggcgtgactg ggcggaggga ctgggcg      47
124 <210> SEQ ID NO: 13
125 <211> LENGTH: 47
126 <212> TYPE: DNA
127 <213> ORGANISM: HIV-1
129 <400> SEQUENCE: 13
130 tcgacgcca gtccctccgc ccagtcacgc ccagtccccc gcccagt      47
132 <210> SEQ ID NO: 14
133 <211> LENGTH: 47

```

## RAW SEQUENCE LISTING

DATE: 05/25/2001

PATENT APPLICATION: US/09/785,632A

TIME: 19:48:04

Input Set : A:\Pto.amc

Output Set: C:\CRF3\05252001\I785632A.raw

```

134 <212> TYPE: DNA
135 <213> ORGANISM: Homo sapiens
137 <400> SEQUENCE: 14
138 ccggagtggg cggtggagtg ggcgtgagtg ggcggatgga gtgggcg      47
140 <210> SEQ ID NO: 15
141 <211> LENGTH: 47
142 <212> TYPE: DNA
143 <213> ORGANISM: Homo sapiens
145 <400> SEQUENCE: 15
146 tcgacgcca ctccatccgc cactcacgc cactccacc gccact      47
148 <210> SEQ ID NO: 16
149 <211> LENGTH: 48
150 <212> TYPE: DNA
151 <213> ORGANISM: Homo sapiens
153 <400> SEQUENCE: 16
154 ccggacatgg gcggagacat gggcgtacat ggcggaaga catgggcg      48
156 <210> SEQ ID NO: 17
157 <211> LENGTH: 48
158 <212> TYPE: DNA
159 <213> ORGANISM: Homo sapiens
161 <400> SEQUENCE: 17
162 tcgacgcca tgtcttcgc ccatgtacgc ccatgtctcc gcccatgt      48
164 <210> SEQ ID NO: 18
165 <211> LENGTH: 120
166 <212> TYPE: DNA
167 <213> ORGANISM: Artificial Sequence
169 <220> FEATURE:
170 <223> OTHER INFORMATION: plasmid sequence
172 <221> NAME/KEY: CDS
173 <222> LOCATION: (1)...(81)
175 <400> SEQUENCE: 18
176 aaa gag ggt ggg tcg acc ttc cgg act ggc cag gaa cgc cca gat ccg      48
177 Lys Glu Gly Gly Ser Thr Phe Arg Thr Gly Gln Glu Arg Pro Asp Pro
178 1      5      10      15
180 cgg gaa ttc aga tct act agt gcg gcc gct aag taagtaagac gtcgagctcg      101
181 Arg Glu Phe Arg Ser Thr Ser Ala Ala Ala Lys
182      20      25
184 ccatcgcggt ggaagcttt      120
186 <210> SEQ ID NO: 19
187 <211> LENGTH: 27
188 <212> TYPE: PRT
189 <213> ORGANISM: Artificial Sequence
191 <220> FEATURE:
192 <223> OTHER INFORMATION: plasmid sequence
194 <400> SEQUENCE: 19
195 Lys Glu Gly Gly Ser Thr Phe Arg Thr Gly Gln Glu Arg Pro Asp Pro
196 1      5      10      15
197 Arg Glu Phe Arg Ser Thr Ser Ala Ala Ala Lys
198      20      25

```

## RAW SEQUENCE LISTING

DATE: 05/25/2001

PATENT APPLICATION: US/09/785,632A

TIME: 19:48:04

Input Set : A:\Pto.amc

Output Set: C:\CRF3\05252001\I785632A.raw

```

200 <210> SEQ ID NO: 20
201 <211> LENGTH: 303
202 <212> TYPE: DNA
203 <213> ORGANISM: Artificial Sequence
205 <220> FEATURE:
206 <223> OTHER INFORMATION: plasmid sequence
208 <221> NAME/KEY: CDS
209 <222> LOCATION: (25)...(291)
211 <400> SEQUENCE: 20
212 gggtcgacct tccggactgg ccag gaa cgc cca tat gct tgc cct gtc gag      51
213                               Glu Arg Pro Tyr Ala Cys Pro Val Glu
214                               1           5
216 tcc tgc gat cgc cgc ttt tct cgc tcg gat gag ctt acc cgc cat atc      99
217 Ser Cys Asp Arg Arg Phe Ser Arg Ser Asp Glu Leu Thr Arg His Ile
218 10           15           20           25
220 cgc atc cac act ggc cag aag ccc ttc cag tgt cga atc tgc atg cgt      147
221 Arg Ile His Thr Gly Gln Lys Pro Phe Gln Cys Arg Ile Cys Met Arg
222           30           35           40
224 aac ttc agt cgt agt gac cac ctt acc acc cac atc cgg acc cac acc      195
225 Asn Phe Ser Arg Ser Asp His Leu Thr Thr His Ile Arg Thr His Thr
226           45           50           55
228 ggc gag aag cct ttt gcc tgt gac att tgt ggg agg aag ttt gcc agg      243
229 Gly Glu Lys Pro Phe Ala Cys Asp Ile Cys Gly Arg Lys Phe Ala Arg
230           60           65           70
232 agt gat gaa cgc aag agg cat acc aaa atc cat tta aga cag aag gat      291
233 Ser Asp Glu Arg Lys Arg His Thr Lys Ile His Leu Arg Gln Lys Asp
234           75           80           85
236 ccgcgggaat cc      303
238 <210> SEQ ID NO: 21
239 <211> LENGTH: 89
240 <212> TYPE: PRT
241 <213> ORGANISM: Artificial Sequence
243 <220> FEATURE:
244 <223> OTHER INFORMATION: plasmid sequence
246 <400> SEQUENCE: 21
247 Glu Arg Pro Tyr Ala Cys Pro Val Glu Ser Cys Asp Arg Arg Phe Ser
248 1           5           10           15
249 Arg Ser Asp Glu Leu Thr Arg His Ile Arg Ile His Thr Gly Gln Lys
250           20           25           30
251 Pro Phe Gln Cys Arg Ile Cys Met Arg Asn Phe Ser Arg Ser Asp His
252           35           40           45
253 Leu Thr Thr His Ile Arg Thr His Thr Gly Glu Lys Pro Phe Ala Cys
254           50           55           60
255 Asp Ile Cys Gly Arg Lys Phe Ala Arg Ser Asp Glu Arg Lys Arg His
256 65           70           75           80
257 Thr Lys Ile His Leu Arg Gln Lys Asp
258           85
260 <210> SEQ ID NO: 22
261 <211> LENGTH: 102

```

## RAW SEQUENCE LISTING

DATE: 05/25/2001

PATENT APPLICATION: US/09/785,632A

TIME: 19:48:04

Input Set : A:\Pto.amc

Output Set: C:\CRF3\05252001\I785632A.raw

```

262 <212> TYPE: DNA
263 <213> ORGANISM: Homo sapiens
265 <220> FEATURE:
266 <221> NAME/KEY: CDS
267 <222> LOCATION: (1)...(102)
269 <400> SEQUENCE: 22
270 acc ggg cag aaa ccg tac aaa tgt aag caa tgt ggg aaa gct ttt gga      48
271 Thr Gly Gln Lys Pro Tyr Lys Cys Lys Gln Cys Gly Lys Ala Phe Gly
272 1          5          10          15
274 tgt ccc tca aac ctt cga agg cat gga agg act cac acc ggc gag aaa      96
275 Cys Pro Ser Asn Leu Arg Arg His Gly Arg Thr His Thr Gly Glu Lys
276          20          25          30
278 ccg cgg      102
279 Pro Arg
283 <210> SEQ ID NO: 23
284 <211> LENGTH: 34
285 <212> TYPE: PRT
286 <213> ORGANISM: Homo sapiens
288 <400> SEQUENCE: 23
289 Thr Gly Gln Lys Pro Tyr Lys Cys Lys Gln Cys Gly Lys Ala Phe Gly
290 1          5          10          15
291 Cys Pro Ser Asn Leu Arg Arg His Gly Arg Thr His Thr Gly Glu Lys
292          20          25          30
293 Pro Arg
296 <210> SEQ ID NO: 24
297 <211> LENGTH: 102
298 <212> TYPE: DNA
299 <213> ORGANISM: Homo sapiens
301 <220> FEATURE:
302 <221> NAME/KEY: CDS
303 <222> LOCATION: (1)...(102)
305 <400> SEQUENCE: 24
306 acc ggg gag aag cca tac aag tgt aag gag tgt ggg aaa gcc ttc aac      48
307 Thr Gly Glu Lys Pro Tyr Lys Cys Lys Glu Cys Gly Lys Ala Phe Asn
308 1          5          10          15
310 cac agc tcc aac ttc aat aaa cac cac aga atc cac acc ggc gaa aag      96
311 His Ser Ser Asn Phe Asn Lys His His Arg Ile His Thr Gly Glu Lys
312          20          25          30
314 ccg cgg      102
315 Pro Arg
319 <210> SEQ ID NO: 25
320 <211> LENGTH: 34
321 <212> TYPE: PRT
322 <213> ORGANISM: Homo sapiens
324 <400> SEQUENCE: 25
325 Thr Gly Glu Lys Pro Tyr Lys Cys Lys Glu Cys Gly Lys Ala Phe Asn
326 1          5          10          15
327 His Ser Ser Asn Phe Asn Lys His His Arg Ile His Thr Gly Glu Lys
328          20          25          30

```

**Please Note:**

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

## VERIFICATION SUMMARY

PATENT APPLICATION: US/09/785,632A

DATE: 05/25/2001

TIME: 19:48:05

Input Set : A:\Pto.amc

Output Set: C:\CRF3\05252001\I785632A.raw

L:1007 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:68  
L:1009 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:68  
L:1033 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:69  
L:1035 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:69  
L:1059 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:70  
L:1061 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:70  
L:1085 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:71  
L:1087 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:71  
L:1115 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:72  
L:1117 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:72  
L:1141 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:73  
L:1143 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:73  
L:1167 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:74  
L:1169 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:74  
L:1197 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:75  
L:1199 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:75  
L:1223 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:76  
L:1225 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:76  
L:1281 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:77  
L:1283 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:77  
L:1305 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:78  
L:1344 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:81  
L:1360 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:82  
L:1383 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:83  
L:1410 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:85  
L:1448 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:88  
L:1463 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:89  
L:1700 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:108  
L:1701 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:108  
L:1716 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:109  
L:1717 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:109  
L:2334 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:150  
L:2336 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:150  
L:2360 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:151  
L:2362 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:151  
L:2386 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:152  
L:2388 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:152  
L:2412 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:153  
L:2414 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:153  
L:2438 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:154  
L:2440 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:154  
L:2464 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:155  
L:2466 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:155  
L:2490 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:156  
L:2492 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:156  
L:2516 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:157  
L:2518 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:157  
L:2542 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:158

## VERIFICATION SUMMARY

PATENT APPLICATION: US/09/785,632A

DATE: 05/25/2001

TIME: 19:48:05

Input Set : A:\Pto.amc

Output Set: C:\CRF3\05252001\I785632A.raw

L:2544 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:158  
L:2568 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:159  
L:2570 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:159  
L:2594 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:160  
L:2596 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:160  
L:2620 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:161  
L:2622 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:161  
L:2646 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:162  
L:2648 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:162  
L:2672 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:163  
L:2674 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:163  
L:2698 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:164  
L:2700 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:164  
L:2724 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:165  
L:2726 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:165  
L:2750 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:166  
L:2752 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:166